UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/886,771	06/21/2001	Paul S. Bradley	15-769 - 4254	6135

22971 7590 11/15/2005

MICROSOFT CORPORATION
ATTN: PATENT GROUP DOCKETING DEPARTMENT
ONE MICROSOFT WAY
REDMOND, , WA 98052-6399

EXAMINER	
MAHMOUDI, HASSAN	
ART UNIT	**PAPER NUMBER**
2165	

DATE MAILED: 11/15/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/886,771	BRADLEY ET AL.	
	Examiner	Art Unit	
	Tony Mahmoudi	2165	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *19 October 2005*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-31* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☐ Claim(s) _____ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *21 June 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08). Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's Request for Continued Examination (RCE) submission and amendments filed on 19-October-2005 has been entered.

Remarks

2. In response to communications filed on 19-October- 2005, claims 1, 8, 18 and 22 have been amended per applicant's request. Claims 1-31 are presently pending in the application, of which, claims 1, 8, 18, and 22 are presented in independent form.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that said subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Fayyad et al (PCT Pub No. WO 99/62007) in view of Kothuri et al (U.S. Patent No. 6,470,344 B1.)

As to claim 1, Fayyad et al. teaches a method for clustering data in a database comprising:

a) providing a database having a number of data records having both discrete and continuous attributes (see page 7, lines 4-6);

b) grouping together data records from the database which have specified discrete attribute configurations (see page 8, lines 5 through page 9, lines 1-13; and see Table 1 and "Cluster Attribute/Value Probability Tables");

c) clustering data records having the same or similar specified discrete attribute configuration based on the continuous attributes to produce an intermediate set of data clusters (see page 11, line 42 through page 12, line 32); and

d) merging together clusters from the intermediate set of data clusters to produce a clustering model (see page 14, lines 26-28; and see figures 8A-8D).

Fayyad et al does not teach performing clustering in two phases, over a discrete attribute and continuous attributes.

Kothuri et al teaches buffering hierarchical index of multi-dimensional data (see Abstract), in which he teaches clustering of data in two phases, over a discrete attribute and continuous attributes (see column 12, lines 40-54, and see column 14, lines 30-65.)

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Fayyad et al by the teaching of Kothuri et al, because including clustering in two phases, over a discrete attribute and continuous

attributes, would enable the system to store different types of data, based on their attributes, into different clusters or groups (e.g. clustering data with attributes having discrete values, determining the number of positive values, and clustering data with attributes having continuous values (range of values), as taught by Kothuri et al (see column 14, lines 30-65.)

As to claims 2, 9, and 23, Fayyad et al. as modified, teaches wherein the clustering model includes a table of probabilities for the discrete data attributes of the data records for a cluster and wherein the cluster model for continuous data attributes comprises a mean and a covariance for each cluster lines (see Fayyad et al, claim 1b).

As to claims 3, 14, and 24, Fayyad et al. as modified, teaches wherein the process of merging of intermediate clusters is ended when a specified number of clusters has been formed (see Fayyad et al, page 8, lines 12-14, where "specified number of clusters" is read on "initial cluster number K=3"; and see claim 14, where "specified number of clusters" is read on "K clusters").

As to claims 4 and 25, Fayyad et al. as modified, teaches wherein the step of merging of intermediate clusters is ended when a distance between intermediate clusters is greater than a specified minimum distance (see Fayyad et al, page 27, line 12 through page 28, line 26, where "distance between intermediate clusters" is read on "stopping criteria" and "specified minimum distance" is read on "the sum of these two numbers" and "the sum of these numbers").

As to claims 5 and 26, Fayyad et al. as modified, teaches wherein the discrete attributes are Boolean and similarity between configurations is based on a distance between bit patterns of the discrete attributes (see Fayyad et al, page 33 where "Boolean" and "bit patterns" is read on "0/1 assignments").

As to claims 6 and 20, Fayyad et al. as modified, teaches wherein one or more of the discrete attributes have more than two possible values and comprising the step of subdividing a discrete attribute having more than two possible values into multiple Boolean value attributes (see Fayyad et al, page 33 where "Boolean" and "two possible values" is read on "0/1 assignments").

As to claims 7 and 27, Fayyad et al. as modified, teaches wherein the step of identifying configurations includes tabulating data records having the same discrete attribute bit pattern and combining the data records from similar configurations before clustering the data records so tabulated based on the continuous attributes (see Fayyad et al, page 33 where "bit pattern" is read on "0/1 assignments).

As to claim 8, Fayyad et al. teaches a method for clustering data in a database comprising:

a) providing a database having a number of data records having both discrete and continuous attributes (see page 14, line 32 through page 15, line 2);

b) counting data records from the database which have the same discrete attribute configuration and identifying a first set of configurations wherein the number of data records of each configuration of the first set of configurations exceeds a threshold number of data records (see page 15, line 21 through page 16, line 15, where "counting data records" is read on "counting the number of data records" and "exceeds a threshold number of data records" is read on "stopping criteria");

c) adding data records from the database not belonging to one of the first set of configurations with a configuration within the first set of configurations to produce a subset of records from the database belonging to configurations in the first set of configurations (see page 15, lines 12-18, where "subset of records" is read on "compressed data"); and

d) clustering the subset of records contained within at least some of the first set of configurations based on the continuous data attributes of records contained within that first set of configurations to produce a clustering model (see page 15, lines 19-27, where "continuous data attributes" is read on "ordered attributes").

For the teaching of "performing a first discrete clustering", and "performing a second continuous clustering", the applicant is directed to the remarks and discussions made in claim 1 above.

As to claim 10, Fayyad et al. as modified, teaches wherein an added record not contained within the first set of configurations is added to one of the first set of configurations based on a distance between a smaller configuration to which the added record belongs during

counting of records in different configurations (see Fayyad et al, page 15, line 24-25, where "counting" is read on "'M' counting").

As to claims 11 and 28, Fayyad et al. as modified, teaches wherein the clustering of records from a configuration based on continuous data attributes results in a variable number of clusters for each configuration based on the number of records in the configuration (see Fayyad et al, page 15, lines 19-32, where "continuous data attributes" is read on "ordered attributes"; and where "variable number of clusters" is read on "scalable clustering process").

As to claim 12, Fayyad et al. as modified, teaches wherein the clustering of records from records falling within a configuration of the first set results in a number of intermediate clusters which are merged together to form the cluster model (see Fayyad et al, page 18, lines 23-31, where "records falling with a configuration" is read on "data points falling within a given cluster").

As to claim 13, Fayyad et al. as modified, teaches wherein intermediate clusters are merged together based on a distance between clusters that is determined based on both continuous and discrete attributes of the intermediate clusters (see Fayyad et al, page 4, line 20 through page 5, line 4, where "clusters are merged" is read on "membership of a given record in a particular cluster"; and see page 19, lines 1-7, where "distance between clusters" is read on "sufficiently 'close' to an existing CS subcluster").

As to claim 15, Fayyad et al. as modified, teaches wherein the merging of intermediate clusters is performed until a distance between two closest clusters is greater than a threshold distance (see Fayyad et al, page 19, line 25 through page 20, line 2).

As to claims 16 and 29, Fayyad et al. as modified, teaches wherein a list of records of each configuration in the first set of configurations is maintained as data records are accessed from the database (see Fayyad et al, page 8, lines 5 through page 9, lines 1-13; and see Table 1 and "Cluster Attribute/Value Probability Tables").

As to claims 17 and 30, Fayyad et al. as modified, teaches where the clustering based on the continuous attributes of records within a configuration is performed using expectation maximization clustering of the continuous attributes (see Fayyad et al, page 4, line 20 through page 5, line 4).

As to claim 18, Fayyad et al. teaches a data processing system comprising:

a) a storage medium for storing a database having a number of data records having both discrete and continuous attributes (see page 7, lines 4-9);

b) a computer for evaluating data records from the database and building a clustering model that describes data in the database (see page 7, lines 1-5); and

c) a database management system including a component for selectively retrieving data records from the database for evaluation by the computer (see page 7, lines 9-11, where "retrieving data records" is read on "brings data from the database");

For the remaining steps of this claim, the applicant is directed to remarks and discussions made in claim 1 above.

As to claim 19, Fayyad et al. as modified, teaches wherein the computer includes a rapid access storage for maintaining a list of data records from the database for data records having a specified discrete attribute configuration to facilitate clustering of the data records based on their continuous attributes (see Fayyad et al, page 5, lines 5-8).

As to claim 21, Fayyad et al. as modified, teaches wherein the rapid access storage of the computer includes a data structure for storing a clustering model (see Fayyad et al, figures 8A-8D).

As to claim 22, Fayyad et al. teaches a computer readable medium containing stored instructions for clustering data in a database comprising instructions for (see page 7, lines 1-11):

a) reading records from a database having a number of data records having both discrete and continuous attributes (see page 7, lines 4-11, where "reading records" is read on "brings data from the database");

For the remaining steps of this claim, the applicant is directed to remarks and discussions made in claims 1 and 8 above.

As to claim 31, Fayyad et al. as modified, teaches where records are assigned to a single cluster during the expectation maximization clustering process (see Fayyad et al, page 4, lines 26-31; and see claim 24).

Response to Arguments

5. Applicant's arguments filed on 19-October-2005 with regard to the common ownership have been fully considered but they are considered moot in view of the new grounds of rejection.

Conclusion

6. Any inquiries concerning this communication or earlier communications from the examiner should be directed to Tony Mahmoudi whose telephone number is (571) 272-4078. The examiner can normally be reached on Mondays-Fridays from 08:00 am to 04:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey Gaffin, can be reached at (571) 272-4146.

tm

November 7, 2005